



Ancient Documents Denoising and Decomposition Using Aujol and Chambolle Algorithm

Mickaël Coustaty, Sloven Dubois, Ménard Michel, Jean-Marc Ogier

► To cite this version:

Mickaël Coustaty, Sloven Dubois, Ménard Michel, Jean-Marc Ogier. Ancient Documents Denoising and Decomposition Using Aujol and Chambolle Algorithm. Kwon and Ogier. Graphics Recognition. New Trends and Challenges, Springer, pp.15-24, 2013, LNCS, 978-3-642-36823-3. 10.1007/978-3-642-36824-0_2 . hal-00916015

HAL Id: hal-00916015

<https://hal.science/hal-00916015>

Submitted on 9 Dec 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Ancient Documents Denoising and decomposition using Aujol and Chambolle Algorithm

Mickael Coustaty¹, Sloven Dubois², Michel Menard¹, and Jean-Marc Ogier¹

¹Laboratoire d'Informatique, Image et Interactions
Université de La Rochelle, Avenue Michel Crepeau
17042 La Rochelle, France
{mcoustat, mmenard, jmogier}@univ-lr.fr

²Université de Lyon, F-42023, Saint-Étienne, France
CNRS, UMR5516, Laboratoire Hubert Curien, F-42000, Saint-Étienne, France
Université de Saint-Étienne, Jean Monnet, F-42000, Saint-Étienne, France
Bâtiment F 18 Rue du Professeur Benoît Lauras, 42000 Saint-Étienne
sloven.dubois@univ-st-etienne.fr

Abstract. With the improvement of printing technology since the 15th century, there is a huge amount of printed documents published and distributed. These documents are degraded by the time and require to be preprocessed before being submitted to image indexing strategy, in order to enhance the quality of images. This paper proposes a new pre-processing that permits to denoise these documents, by using a Aujol and Chambolle algorithm. Aujol and Chambolle algorithm allows to extract meaningful components from image. In this case, we can extract shapes, textures and noise. Some examples of specific processings applied on each layer are illustrated in this paper.

Keywords: Historical documents, decomposition, denoising, Aujol and Chambolle

1 Introduction

With the improvement of printing technology since the 15th century, there is a huge amount of printed documents published and distributed. At this period, the printed book quickly becomes a regular object in the world. By 1501, there were 1000 printings shops in Europe, which were at the basis of the production of 35,000 titles and 20 million copies. Since that time, most of these books have been falling into decay and degrading. With the disappearing of these books, the knowledge of our ancestors is also vanishing. Therefore, there are a lot of attempts to keep, organize and restore ancient printed documents. With the better digital technology, one of the preservation methods of these old documents is digitizing. This explains the large number of digitizing campaigns that one can find all over the world. However, digitizing documents is not beneficial without the ability to retrieve and extract the information from them which could be done by using techniques of document analysis and recognition. The extraction of this kind of information requires the use of high level features extraction techniques, which can be strongly altered if images are degraded. This paper proposes a set of contributions related to the use of Aujol and Chambolle algorithm for the enhancement of the quality of images.

2 NaviDoMass and Historical document challenges

2.1 NaviDoMass

NAVIDOMASS (Navigation into Documents Masses) is a french collaborative project. With the collaboration of seven laboratories in France, the global objective of this project is to build a framework to derive benefit from historical documents. It aims at preserving and providing public accessibility to this national heritage. NAVIDOMASS was a three-year project completed in 2010. It was established on four principles: anywhere (global access), anyone (public and multilingual), anytime and any media (accessible through various channels such as world wide web, smartphone, etc.). The focus of NAVIDOMASS was on five studies: (1) user requirement, participative design and ground truthing, (2) document layout analysis and structure based indexing, (3) information spotting, (4) structuring the feature space [8, 9] and (5) interactive extraction and relevance feedback.

As a part of NAVIDOMASS project, this paper focuses on the preprocessing part, or enhancing the graphical part of the document in order to use it in indexing. However, the main interest of this study is based on specific graphics called drop caps and is inspired by [14] and [16]. Even if the use case of this paper is reduced to drop caps analysis (also called dropcaps or ornamental letters), the aim of this study is to develop complex graphic pre-processing techniques. By complex graphic, we mean here any graphic document which does not respect a rigorous structure, as one can find in the domain of symbol recognition for instance.

2.2 Historical documents specificities

Historical documents are specific in many different points. The first difference between historical and actual documents relies on the structure and the variation of structure of these documents. Actually, when you buy a newspaper, you can usually find texts, graphics and images. All these elements, and their spatial organization, compose the structure of the document. In contemporary documents, this structure is generally well-defined, can be clearly identified, and is the same for a category of document (all books have a title, a subtitle, ...). In the case of heritage documents, the structure was specific to each document and to each printer. An example of this heterogeneity can be observed in Figure 1, where documents from different libraries are presented.

The second difference relies on the method used to print these documents. At the beginning of the print technology, a machine applied a wood stamp on paper to deposit ink. These stamps degrades themselves with time and the number of impressions. The usury of stamps deteriorate the quality of books, texts and illustrations which create a big quantity of variations for a same model. For example, one can see in Figure 1 that the same image differs in different books.

From these two problems, we can see that historical documents are hard to segment and to retrieve in database. The variations of structure prevent to create a unique treatment to compare documents and their contents. Moreover, these documents are rich in term of information and complexity and identifying an element only by a global description of the document is impossible.



Fig. 1. Different types of documents and variations of an image

Finally, heritage documents are subjected to the ageing, and all the marks of time (mildew, thinning of paper, ...). All these problems create a non uniform background in which we can find the text of the back page, variation of gray-levels and different elements which make though the extraction and the segmentation of content of documents. In this next of this paper, we propose a new method, based on the use of the Aujol and Chambolle algorithm, to denoise images of documents and to simplify them by decomposing each image in different layers of information. We first work on a treatment adapted to image containing principally text and in a second time images of drop caps.

3 Grayscale decomposition model of J-F Aujol and A. Chambolle

Decomposing an image into meaningful components appears as one of major aims in recent development in image processing. This kind of model comes from well known ill-posed problem in image processing, the aim of which is to recover an ideal image u , from a degraded observation f as follow: $f = Au + v$ (with A a linear operator representing blur and v the noise, often additive). Retrieving the original image can then be seen as a deconvolution which is an ill-posed problem: operator A is not boundedly invertible in $H = L^2(a, b)$ (A is compact in H).

The first goal was image restoration and denoising, but following the ideas of Yves Meyer [11], in total variation minimization framework of L. Rudin, S. Osher and E. Fatemi [10], image decomposition into geometrical and oscillatory (i.e texture) component appears as useful and very interesting way for our image analysis case. Indeed, we want to obtain the main structure of images in order to properly segment drop caps, independently of textured parts and avoid acquisition problems like noise.

Our aim is to catch the pure geometrical component in an image independently of texture and noise. For this, in [2], J-F. Aujol and A. Chambolle propose a decomposition model which splits a grayscale image into three components: the first one, $u \in BV$

¹, containing the structure of the image, a second one, $v \in G$ ², the texture, and the third one, $w \in E$ ³, the noise. For better comprehension of different spaces, see [1–3, 7]. One of the main advantage of this decomposition is its reversibility, as when an image has been decomposed, we are able to reconstruct it, with a precision related to the rounding errors.

This decomposition model is based on a minimization of this discretized functional F :

$$\inf_{(u,v,w) \in X^3} F(u, v, w) = J(u) + J^* \left(\frac{v}{\mu} \right) + B^* \left(\frac{w}{\lambda} \right) + \frac{1}{2\alpha} \|f - u - v - w\|_{L^2}^2$$

where $J(u)$ is the total variation related to the extraction of the geometrical component, $J^* \left(\frac{v}{\mu} \right)$, $B^* \left(\frac{w}{\lambda} \right)$ are the Legendre-Fenchel transforms⁴ of respectively J and B [3] for the extraction of texture and noise components, parameter α controls the L^2 - norm of the residual $f - u - v - w$ and X is the discrete euclidean space $\mathbb{R}^{N \times N}$ for images of size $N \times N$.

For minimizing this fonctionnal, Chambolle's projection algorithm is used [2]. The Chambolle's projection P on space λB_G ⁵ of f is denoted $P_{\lambda B_G}(f)$ and is solved by an iterative algorithm. This algorithm starts with $P^0 = 0$ and for each pixel (i, j) and at each step $n + 1$ we have:

$$P_{i,j}^{n+1} = \frac{P_{i,j}^n + \tau \left(\Delta \operatorname{div} (P^n) - \frac{f}{\lambda} \right)_{i,j}}{1 + \tau \left| \Delta \operatorname{div} (P^n) - \frac{f}{\lambda} \right|_{i,j}} \quad (1)$$

¹ $BV(\Omega)$ is the subspace of functions $u \in L^1(\Omega)$ such that the following quantity, called the total variation of u , is finite:

$$J(u) = \sup \left\{ \int_{\Omega} u(x) \operatorname{div}(\xi(x)) dx \right\}$$

such that $\xi \in C_c^1(\Omega, \mathbb{R}^2)$, $\|\xi\|_{L^\infty(\Omega)} \leq 1$

² G is the subspace introduced by Meyer for oscillating patterns. G denotes the Banach space composed of the distributions f which can be written $f = \partial_1 g_1 + \partial_2 g_2 = \operatorname{div}(g)$ with g_1 and g_2 in $L^\infty(\Omega)$. On G , the following norm is associated:

$$\|v\|_G = \inf \{ \|g\|_{L^\infty(\Omega, \mathbb{R}^2)} / v = \operatorname{div}(g), \\ g = (g_1, g_2), |g(x)| = \sqrt{(|g_1|^2 + |g_2|^2)(x)} \}$$

³ E is another dual space to model oscillating patterns. $\dot{B}_{1,1}^1$ is the usual homogeneous Besov space and the dual space of $\dot{B}_{1,1}^1$ is the Banach space $E = \dot{B}_{-1,\infty}^\infty$

⁴ The Legendre-Fenchel transform of F is given by $F^*(v) = \sup_u (\langle u, v \rangle_{L^2} - F(u))$, where $\langle \cdot, \cdot \rangle_{L^2}$ stands for the L^2 inner product [15]

⁵ $\lambda B_G = \{f \in G / \|f\|_G \leq \lambda\}$

In [4] a sufficient condition ensuring the convergence of this algorithm is given: $\tau \leq \frac{1}{8}$. To solve (3), the authors propose the algorithm of figure 2.

Algorithm :

Initialization: $u = v = 0$

1. u and v have been previously computed, we estimate: $\tilde{w} = P_{\delta B_E}(f - u - v)$
2. then we compute: $\tilde{v} = P_{\mu B_G}(f - u - \tilde{w})$
3. and we finally obtain: $\tilde{u} = f - u - \tilde{v} - \tilde{w} - P_{\lambda B_G}(f - \tilde{v} - \tilde{w})$

This operation is repeated until :

$$\max(|\tilde{u} - u|, |\tilde{v} - v|, |\tilde{w} - w|) \leq \varepsilon$$

Fig. 2. Grayscale image decomposition algorithm

In [2], the authors replace $P_{\delta B_E}(f - u - v)$ by $f - u - v - W_{ST}(f - u - v, \delta)$ where $W_{ST}(f - u - v, \delta)$ stands for the wavelet soft-thresholding of $f - u - v$ with threshold δ defined by :

$$S_{\delta}(d_i^j) = \begin{cases} d_i^j - \delta \text{sign}(d_i^j) & \text{if } |d_i^j| > \delta \\ 0 & \text{if } |d_i^j| \leq \delta \end{cases} \quad (2)$$

where d_i^j is the wavelet coefficient, j the resolution and $i \in \{x, y, xy\}$.

The figure 3 shows the application of grayscale decomposition model of J-F Aujol and A. Chambolle into an image.

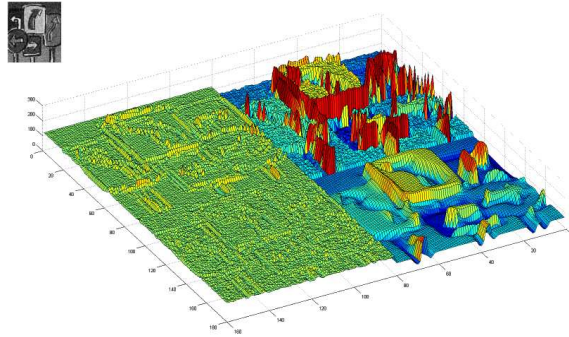


Fig. 3. Image decomposition. From top to bottom and left to right: Texture component (v), original noisy image (f), noise component (w) and regular or geometrical part (u).

In the next section, the benefits of this algorithm on historical documents are presented.

4 Pre-processing for historical documents based on the Aujol and Chambolle algorithm

In all this section , we will call "shape layer" the image associated to the regular and geometrical part, and "texture layer" the texture component, obtained by the Aujol and Chambolle algorithm.

4.1 Denoising Text

We can see in the Figure 4(a) that images of old documents are degraded and some see through text of the back page appear. The Aujol and Chambolle algorithm is used on old documents images to separate informations in three layers. The two first layers (shapes and texture layers) are exposed in Figure 4(b) and in Figure 4(c).

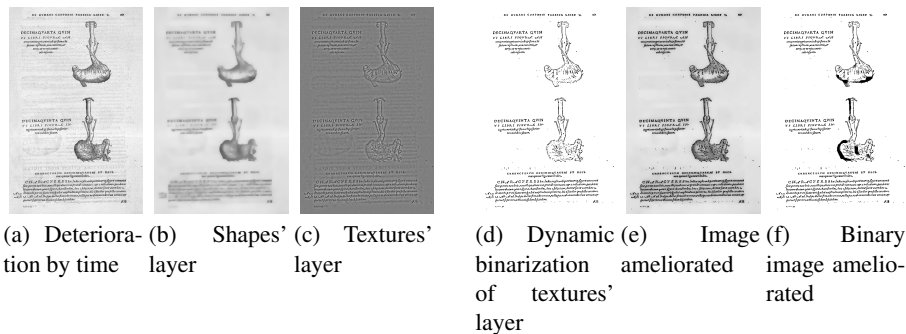


Fig. 4. A page degraded by time. The three layers extract with the Aujol and Chambolle algorithm (a, b, c) and image ameliorated for segmentation (d, e, f)

One can see that the Aujol and Chambolle algorithm allows to distinctly separate shapes from textures of the original image. In our specific case of literal and graphic document, the layer associated to texture contains all the text of the back page while the shape layer permits to reduce effect of time on the background. By reducing the number of colours to two, one can easily isolates textures of the front from the back one. Once we get the texture of the front page, we add it to the shape layer of the original image to obtain the image in Figure 4(e). And finally, a classical binarization permits to obtain the document presented in Figure 4(f) which contains all the information of the original document with text and graphics. The new version of the document is easier to segment in three parts : text / graphics / background and text can be recognized directly.

4.2 Simplifying Drop Caps

Images composed only of graphics, like drop caps, are very rich in information. One can see in Figure 5 that Aujol and Chambolle algorithm simplifies the image and permits to obtain three layers of information on which we can apply specific preprocessings. For example, shapes' layer permits a better and easier shapes' extraction for recognition. In addition, shapes obtained are cleaned, allowing us to abstain from the noise, what offers greater robustness for segmentation and description.

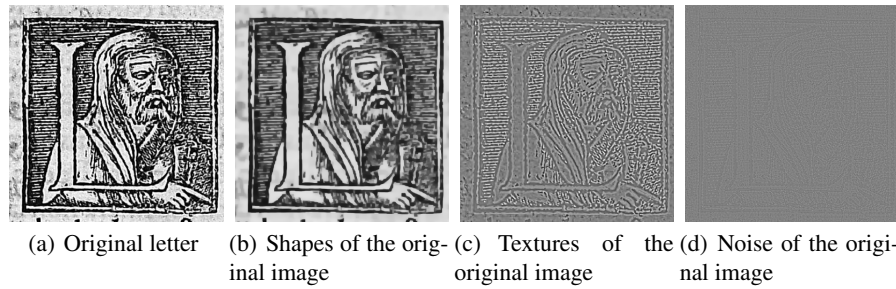


Fig. 5. An example of drop cap using Aujol and Chambolle algorithm

This decomposition gives us simplified images and particularly we have worked on shapes' layer to extract most significant shapes like the letter or faces [6], arms, legs [5], etc. All these elements are useful for historians to retrieve similar images by content. Another application we have developed consisted in applying a segmentation process on textures to extract relevant ones which correspond to important elements, like in [12].

Furthermore, we can see that fundamental elements of drop caps are composed of one or many shapes, mixed with one or many textured regions. A global approach, the aim of which is to describe an image without dissociating these two kind of information, can't precisely describe this kind of elements and these images (main features are dedicated to one kind of information). It becomes then necessary to separate shapes from textures, by simplifying the image.

One solution, to simplify images, consists in using a decomposition that allows separating different kind of elements in the image in different layers. Each layer is then composed of elements with the same properties (only textured regions, only shapes, ...). It is then possible to apply classical features computation adapted to the content of the layer to get a specific description of the layer content. Finally, the global description of the image relies on the combination of all these descriptions.

4.3 Evaluation

The evaluation of such a system is a fundamental point because it guarantees its possible usability by the users, and because it permits to have an objective regard on the

system. In the context of such a project, the implementation of an objective evaluation device is quite difficult, because of the variability of the user requirements on the one hand (historians researchers, netsurfers, are likely to retrieve many different information which can be very different the one from the others), and because this filtering stage in a preliminary process in the global indexing system. However it appears to be important to evaluate it.

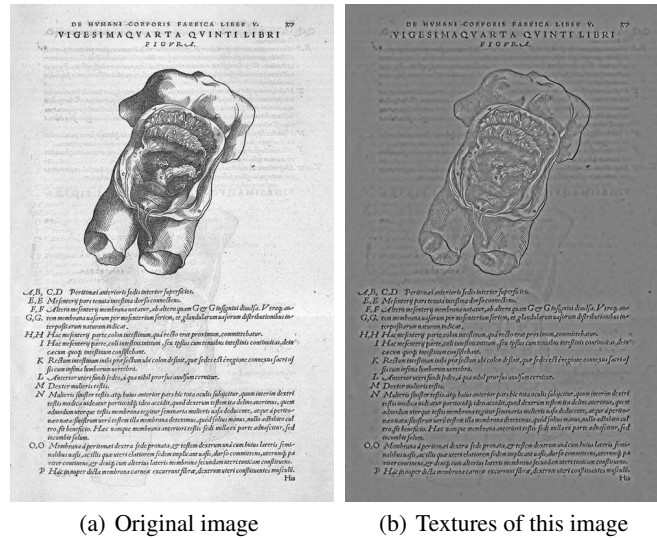


Fig. 6. An example of surimpressing

In the context of NAVIDOMASS project, considering the fundamental aspect of this filtering stage, we propose to use signal processing evaluation criteria: the Signal Noise Ratio (SNR). In the context of such images databases, the main difficulty is to implement a groundtruth for the evaluation of a processing. Considering this difficulty, we decided to use a synthetic noise generator applied on a set of cleaned images, and to evaluate the capacity of our system to remove this noise.

The protocol that we have implemented is the following one:

1. We considered a significant set of images of newspaper documents (see Figure 7). We chose images of newspaper because they have an homogenous background, a good contrast between text and background and no initial noise
2. We generated some additive synthetic noise on these images by using different parameters for this additive noise
3. On each of these images, we applied Aujol and Chambolle algorithm, in order to separate the different layers: shape, texture and noise
4. A binarization stage is applied on the texture layer in order to separate the texture corresponding to the text layer, and the texture coming from issuing from surim-

pressing. As one can see on Figure 6, thanks to this dynamic binarization stage ([13]), all the layer issuing from the background character layer is removed

5. A combination XOR of the texture layer remaining at the issue of the stage 4, and the shape layer is operated in order to build the filtered image
6. At the issue of this stage 5, the resulting image is the filtered image. A computation of its SNR is operated and compared to the original image one

As one can see, this protocol has been applied on a set of document images, on which we have applied different additive synthetic noises, with different statistics. With this



Fig. 7. An example of newspaper

protocol, more than 86% of the 29 images of newspapers which have been tested gave a better SNR. The results are very satisfying, since, the removal of the noise is very efficient (evaluation based on the SNR). These excellent results are validated and confirmed by our historian experts, in terms of pre-processing for further indexing operations.

5 Conclusion and Perspectives

This paper present a new method to simplify and denoise old documents images in order to index them. This process relies on a Aujol and Chambolle algorithm which allow to obtain three layers of information: 1) regularized one which contain shapes, 2) textures one which contain all the oscillating elements and 3) noise one which correspond to the original image less the two first layers. From this decomposition, one can apply specific treatments on one or many layers. In our case, we used shapes' and textures' layers in two example: first one to pick out text of back page and a second one to extract region of interest. Experimentations are promising and we will ameliorate these results

in future works. One interesting perspective that we plan to study consists in conjointly analysing the spatial organisation of the shapes regions against the textures ones. This analysis will permit to take into account the semantic of these complex images.

References

1. Aujol, J.F., Aubert, G., Feraud, L.B., Chambolle, A.: Image decomposition into a bounded variation component and an oscillating component. *Journal of Mathematical Imaging and Vision* 22(1), 71–88 (Jan 2005), <http://dx.doi.org/10.1007/s10851-005-4783-8>
2. Aujol, J.F., Chambolle, A.: Dual norms and image decomposition models. *International Journal of Computer Vision* 63(1), 85–104 (2005), <http://dx.doi.org/10.1007/s11263-005-4948-3>
3. Aujol, J.F., Gilboa, G., Chan, T., Osher, S.: Structure-texture image decomposition - modeling, algorithms, and parameter selection. *International Journal of Computer Vision* 67(1), 111–136 (2006), <http://dx.doi.org/10.1007/s11263-006-4331-z>
4. Chambolle, A.: Total variation minimization and a class of binary mrf models. *EMMCVPR 3757 of Lecture Notes in Computer Sciences*, 136–152 (2005)
5. Coustaty, M., Bouju, A., Bertet, K., Louis, G.: Using ontologies to reduce the semantic gap between historians and image processing algorithms. In: *IEEE International Conference on Document Analysis and Recognition*, Beijing, China. pp. 156–160 (2011)
6. Coustaty, M., Pareti, R., Vincent, N., Ogier, J.M.: Towards historical document indexing: extraction of drop cap letters. *IJDAR* 14(3), 243–254 (2011)
7. ELHAMIDI, A., MENARD, M., LUGIEZ, M., GHANNAM, C.: Weighted and extended total variation for image restoration and decomposition. *Pattern Recognition* 43(4), 1564–1576 (2010)
8. H.Chouaib, S.Tabbone, O.Ramos, Cloppet, F., N.Vincent: Feature selection combining genetic algorithm and adaboost classifiers. In: *ICPR'08. Florida* (2008)
9. Jouili, S., Tabbone, S.: Applications des graphes en traitement d'images. In: *ROGICS'08*. pp. 434–442. University of Ottawa, Canada and University of Sfax, Tunisia, Mahdia Tunisia (2008)
10. L.Rudin, S.Osher, E.Fatemi: Nonlinear total variation based noise removal. *Physica D* 60, 259–269 (1992)
11. Meyer, Y.: Oscillating patterns in image processing and nonlinear evolution equations. *The fifteenth dean jacqueline B. Lewis Memorial Lectures* (2001)
12. Nguyen, T.T.H., Coustaty, M., Ogier, J.M.: Bags of strokes based approach for classification and indexing of drop caps. In: *IEEE International Conference on Document Analysis and Recognition*, Beijing, China. pp. 349–353 (2011)
13. Otsu, N.: A threshold selection method from grey scale histogram. *IEEE Trans. on Syst. Man and Cyber.* (1979)
14. Pareti, R., Vincent, N.: Ancient initial letters indexing. In: *ICPR '06: Proceedings of the 18th International Conference on Pattern Recognition*. pp. 756–759. IEEE Computer Society, Washington, DC, USA (2006)
15. Starck, J.L., ELad, M., Donoho, D.: Image decomposition via the combination of sparse representation and variational approach. *IEEE Trans. Image Process* (2005)
16. Uttama, S., Loonis, P., Delalandre, M., Ogier, J.M.: Segmentation and retrieval of ancient graphic documents. In: *GREC*. pp. 88–98 (2005)